Arya Lokesh Gowda

+1 (214)972-5529 | itsaryalg@gmail.com | https://www.linkedin.com/in/arya-l-gowda | https://aryaalg11.github.io/portfolio/

SKILLS

Data: PostgreSQL, NoSQL, DBeaver, MySQL, Oracle SQL, MongoDB, Salesforce, Snowflake, Netezza, Hadoop, MS Excel (VLOOKUP, Pivot Tables, Macros), Salesforce, SAS, Apache Spark, Hadoop, Hive, Jupyter.
 data cleaning, data preprocessing, exploratory data analytics, data visualization
Business Intelligence: Tableau, Power BI, MS Access
Programming Languages: Python, SQL, Java, Shell Scripting, R, C, C++, MATLAB
Cloud : Amazon Web Services (EC2, Redshift, Lambda, Glue, Athena, Sagemaker, IAM), Google Cloud Platform
ETL and Workflow: Airflow, JIRA, ControlM, GitHub, GitLab, Talend, Jenkins, Informatica, Kafka

EXPERIENCE

Intern- Data Analyst-I | *Nelnet* | Nebraska, USA   *Aug 2024 – Present*
- Formulated financial models and analyzed KPIs to enhance strategic decision-making, improving forecast accuracy by 5% and driving operational efficiency gains of 8%
- Automated financial reporting processes such as Profitability , Revenue reports, etc., enabling faster decision-making for stakeholders and integrated diverse systems to optimize ledger management

Data Engineer | *Infosys* | Bangalore, India   *Oct 2021 – Aug 2023*
- Architected a scalable data warehouse handling 10M+ daily records, with robust data governance protocols that reduced data quality issues by 40%, resulting in a $150K (3%) reduction in operational costs
- Orchestrated end-to-end ETL pipelines processing 500GB+ daily data across 12 source systems, reducing processing time from 6 hours to 45 minutes
- Led development of ML models for customer segmentation and predictive analytics, improving accuracy by 30% and generating $122K additional revenue

Data Analytics Intern | *Entuple Solutions* | Bangalore, India   *Jun 2020 – Jan 2021*
- Deployed machine learning models for product optimization, resulting in 15% performance improvement and $50K annual cost savings across manufacturing operations
- Crafted dashboards visualizing key metrics for 10+ stakeholders, automating reports that saved time by 75%

PROJECTS

Covid-19 Vaccine Sentiment Analysis | *Twitter API, VADER, Time Series Analysis*   *Jul 2024-Aug 2024*
- Engineered scalable pipeline processing 100K+ tweets daily, implementing custom NLP preprocessing that improved accuracy to 85% across 5 major vaccine types
- Developed dashboards correlating sentiment trends with 50+ major vaccine events, identifying 3 key message types that increased vaccine acceptance by 12% in target demographics

Instagram Reach Analysis | *Passive Aggressive Regressor, Seaborn, Statsmodels*   *Jun 2024-July 2024*
- Built ensemble ML model analyzing 20+ engagement metrics across 10,000 posts, achieving 92.8% accuracy and identifying optimal posting windows that increased reach by 35%
- Implemented A/B testing framework for content strategy, increasing average engagement rate to 14% and driving 25% growth in follower acquisition

Fetal Health Classification System | *PySpark, scikit-lean, MLflow, Autoencoders*   *Jan 2024-April 2024*
- Designed a distributed ML pipeline processing 5000+ medical records daily, achieving 97% sensitivity in critical case detection, potentially impacting 1000+ patient outcomes annually
- Engineered custom Autoencoder architecture for anomaly detection, reducing false positives by 40% and achieving 99% precision in high-risk pregnancy identification

Customer Data Stratification | *NumPy, Matplotlib, TensorFlow*   *Sep 2023-Nov 2023*
- Led 5-person team in developing RFM segmentation model analyzing 50,000+ customer profiles, generating $70,000 additional revenue through campaigns targeting 4 distinct customer segments
- Refined and deployed BERT-based NLP model processing 100,000+ customer interactions, achieving 96% accuracy in behavior prediction and driving 13% sales increase

EDUCATION

Northeastern University | *Boston, MA*   *Spring 2025*
 Candidate for M.S in Data Analytics Engineering   *GPA 3.85*
 Focusing: Data Mining, Data Warehousing and Data Management in Analytics, Computation and Visualizations
Visvesvaraya Technological University | *Bangalore, India*   *Jul 2021*
 B.E in Electronics and Communication Engineering
 Focusing: Data Structures in C and JAVA, Linear Algebra, Information Theory and Algorithms